



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:

Dirk Labudde et al.

Title: DEVICE AND PROCESS FOR THE ASSIGNMENT OF NMR SIGNALS OF POLYPEPTIDES

Docket No.: 1708.005US1

Serial No.: 10/799447

Filed: March 11, 2004

Due Date: N/A

Examiner: Unknown

Group Art Unit: Unknown

MS Non-Fee Amendment

Commissioner for Patents

P.O. Box 1450

Alexandria, VA 22313-1450

We are transmitting herewith the following attached items (as indicated with an "X"):

A return postcard.
 Communication Regarding Filing of Priority Document (1 Page).
 Priority Document - German Application No.: 101 44 661.6 (37 pgs.).

If not provided for in a separate paper filed herewith, Please consider this a PETITION FOR EXTENSION OF TIME for sufficient number of months to enter these papers and please charge any additional fees or credit overpayment to Deposit Account No. 19-0743..

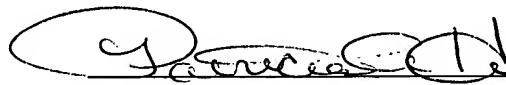
SCHWEGMAN, LUNDBERG, WOESSNER & KLUTH, P.A.
Customer Number 21186

By: 
Atty: William F. Prout
Reg. No. 33,995

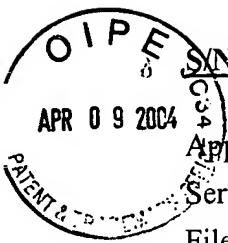
CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail, in an envelope addressed to: MS Non-Fee Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on this 10th day of April, 2004.

PATRICIA A. HULTMAN

Name

Signature 

SCHWEGMAN, LUNDBERG, WOESSNER & KLUTH, P.A.
(GENERAL)



S/N 10/799447

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Dirk Labudde et al. Examiner: Unknown
Serial No.: 10/799447 Group Art Unit: Unknown
Filed: March 11, 2004 Docket: 1708.005US1
Title: DEVICE AND PROCESS FOR THE ASSIGNMENT OF NMR SIGNALS OF
POLYPEPTIDES

COMMUNICATION REGARDING FILING OF
PRIORITY DOCUMENT

Commissioner for Patents
P.O.Box 1450
Alexandria, VA 22313-1450

In accordance with the requirements for claiming right of priority under 35 U.S.C. 119, enclosed for filing in connection with the above-identified application is a certified copy of Applicant's prior application, German Application No. 101 44 661.6, filed September 11, 2001.

Conclusion

If necessary, please charge any additional fees or credit overpayment to Deposit Account No. 19-0743.

Respectfully submitted,

DIRK LABUDDE ET AL.

By their Representatives,

SCHWEGMAN, LUNDBERG, WOESSNER & KLUTH, P.A.
P.O. Box 2938
Minneapolis, MN 55402
(612) 373-6968

Date April 6, 2004

By

William F. Prout

Reg. No. 33,995

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail, in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on this 6 day of April, 2004.

Patricia A. Ultman
Name

Signature

BUNDESREPUBLIK DEUTSCHLAND



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 101 44 661.6

Anmeldetag: 11. September 2001

Anmelder/Inhaber: Forschungsverbund Berlin eV, 12489 Berlin/DE

Bezeichnung: Vorrichtung und Verfahren zur Zuordnung der NMR-Signale von Polypeptiden

IPC: G 01 N 24/08

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 19. März 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

A handwritten signature in black ink, appearing to read "Ebert", is placed over the typed name of the President.

Ebert

Anmelder:

5

Forschungsverbund Berlin e.V.
Rudower Chaussee 17
D-12489 Berlin / DE

10

Vorrichtung und Verfahren zur Zuordnung der NMR-Signale von Polypeptiden

15

Die Erfindung betrifft ein Analysesystem zur automatisierten Analyse eines Satzes von NMR-Spektren, welcher für eine n Aminosäuren umfassende Polypeptidkette aufgezeichnet wurde
20 sowie ein Verfahren zu automatisierten Analyse eines Satzes von NMR-Spektren.

Die NMR-Spektroskopie hat sich in den vergangenen Jahren als Methode für die Strukturaufklärung von kleinen Proteinen und DNA-Fragmenten etabliert. Sie erlaubt die Untersuchung von biologischen Makromolekülen in Lösung – unter besonderer Berücksichtigung
25 dynamischer Phänomene – und stellt damit eine zur Röntgenkristallographie komplementäre Methode dar.

Die NMR-spektroskopische Untersuchung von Proteinen blieb zunächst – mit wenigen Ausnahmen – auf relativ kleine Vertreter mit einer Größe von bis zu 80 Aminosäuren beschränkt,
30 da sich für größere Proteine Signalüberlagerungen in den 2D-Spektren als limitierend erwiesen. Erst die Einführung von dreidimensionalen und vierdimensionalen NMR-Techniken (3D- und 4D-NMR) ermöglichte die Überwindung dieser Barriere. In Verbindung mit einer Markierung der Proteine mit ²H, ¹³C und ¹⁵N lassen sich heute Systeme mit einem Molekulargewicht von bis zu 50 kD untersuchen. Die Größe der noch untersuchbaren Proteine wird im
35 wesentlichen durch die mit zunehmender Masse kürzer werdende transversale Relaxationszeit bestimmt.

Die Vielzahl der existierenden mehrdimensionalen NMR-Techniken werden im Rahmen von Strukturbestimmungsprojekten für zwei unterschiedliche Teilschritte benötigt. In einem ersten Schritt müssen alle ^1H , ^{13}C und ^{15}N -Signale eines Proteins zugeordnet werden. Bei diesem Zuordnungsschritt muss zu diesen magnetisch aktiven Kernen im Protein das entsprechende Signal im Spektrum gefunden werden. Für diese Aufgabe der Zuordnung gibt es spezielle Pulssequenzen. In dem Artikel „Protein Structure Determination with Three- and Four-Dimensional NMR Spectroscopy“ von H. Oschkinat et al., Angew. Chem. Int. Ed. Engl. 1994, 33, Seiten 277-293 wird ein Überblick über die verschiedenen zur Proteinstrukturbestimmung verwendeten Experimente und Pulssequenzen gegeben.

10

In dem Artikel „MUSIC, Selective Pulses, and Tuned Delays: Amino Acid Type-Selective ^1H / ^{15}N Correlations, II“ von M. Schubert et al., Journal of Magnetic Resonance 148, 2001, Seiten 61-72 sind eine Anzahl von aminosäuretyp-spezifischen ^1H / ^{15}N -Experimenten beschrieben, in denen unter Ausnutzung der Seitenkettentopologie die Signale eines bestimmten Typs von Aminosäure (z. B. Ser) oder einer bestimmten Gruppe von Aminosäuren (z. B. Ile/Val) enthalten sind. Die zur Durchführung dieser aminosäuretyp-spezifischen 2D-Experimente benötigten Pulssequenzen lassen sich auf einfache Weise aus den zur Bestimmung der Hauptkettenstruktur verwendeten Tripelresonanz-Experimenten ableiten.

20

Nachdem die Zuordnung vollständig durchgeführt ist, können mit Hilfe anderer NMR-Techniken Strukturparameter des Proteins gesammelt werden. Dieser zweite Schritt baut auf der im ersten Schritt erhaltenen Zuordnung auf. Beispielsweise lassen sich mit Hilfe der verschiedenen mehrdimensionalen Versionen des NOESY-Experiments Abstände zwischen verschiedenen magnetisch aktiven Kernen ermitteln. Die so erhaltenen Strukturparameter dienen als Eingangsgrößen für Strukturbestimmungs-Softwarepakete. Derartige Strukturberechnungs-Programme erzeugen aus den zugeführten Strukturinformationen ein dreidimensionales Modell des Polypeptids.

30

Gegenwärtig werden die verschiedenen Schritte der Proteinstrukturanalyse mit semiautomatischen Prozeduren und meist hauseigener Software in den verschiedenen NMR-Forschungsgruppen durchgeführt. Die vielen Versuche, insbesondere den Prozeß der Zuordnung zu erleichtern, haben unter anderem zu sog. elektronischen Zeichentischen geführt, in denen die Spektren auf dem Bildschirm dargestellt und mit vom Programm zur Verfügung

gestellten Hilfsmitteln, wie automatischem Peakpicking und der Möglichkeit zur Spektrenüberlagerung, zugeordnet werden.

Viele Verfahren zur automatischen Zuordnung von NMR-Signalen basieren auf der Verwendung von Kreuzsignallisten, mit denen die Frequenzkoordinaten der Kreuzsignale erfasst werden. Diese Kreuzsignallisten lassen sich mit Hilfe von kombinatorischen Ansätzen auswerten, die Vergleiche zwischen den in den Kreuzsignallisten enthaltenen Frequenzen vorsehen. Wenn die Zuordnung mit Hilfe von Kreuzsignallisten vorgenommen wird, muss eine Reihe von Nachteilen in Kauf genommen werden. So treten bei Spektren mit niedrigem Signal-Rausch-Verhältnis oder starkem T1-Rauschen spektrale Artefakte auf, welche unerwünschte zusätzliche Einträge in die Kreuzsignallisten verursachen und eine erfolgreiche Anwendung der kombinatorischen Verfahren erschweren. Wenn die Kreuzsignale des Spektrums sehr dicht liegen, können die einzelnen Kreuzsignale nicht mehr gegeneinander aufgelöst werden. In diesem Fall ergibt sich eine Kreuzsignalliste, in der verschiedene Einträge fehlen oder verfälscht sind.

Aus diesen Gründen ist man von der Verwendung von Kreuzsignallisten abgekommen und versucht statt dessen, die in den NMR-Spektren enthaltenen Signalmuster mit Hilfe von alternativen Verfahren zu erfassen. Ein derartiger Ansatz ist in dem Artikel „Tools for the automated assignment of high-resolution three-dimensional protein NMR spectra based on pattern recognition techniques“ von D. Croft et al., Journal of Biomolecular NMR, 10, 1997, Seiten 207-219 beschrieben. In diesem Artikel wird speziell auf die Signalmuster-Erkennungssoftware CATCH23 eingegangen. Diese Software verwendet Suchmasken zur Analyse der NMR-Spektren und führt mit Hilfe einer Kombination von Suchmasken eine Mustersuche durch. Eine solche Kreuzsignalmuster-Suchmaske erfasst eine Vielzahl von Suchbereichen für die zu erwartenden Kreuzsignale eines bestimmten Hauptketten- bzw. Seitenkettenfragments. Beispielsweise können mit einer Kreuzsignalmuster-Suchmaske sämtliche von der Aminosäure Threonin verursachten Kreuzsignale erfasst werden. Falls die Kreuzsignalmuster-Suchmaske die entsprechenden Peaks identifiziert, kann eine Zuordnung zwischen diesen Peaks und der Aminosäure Threonin vorgenommen werden.

Häufig existieren jedoch mehrere Möglichkeiten, eine Zuordnung zwischen den Kreuzsignalen einerseits und der Molekülstruktur andererseits zu treffen. Diese Vieldeutigkeit bei der

Zuordnung macht oftmals ein manuelles Eingreifen erforderlich. Die in der publizierten Version der Software CATCH23 definierten Kreuzsignalmuster-Suchmasken ermöglicht noch keine hinreichende Stabilität und Sicherheit für eine vollautomatische Zuordnung der Signal-peaks.

5

Es ist daher Aufgabe der Erfindung, eine Vorrichtung sowie ein Verfahren zur automatisierten Zuordnung der NMR-Signale eines Satzes von NMR-Spektren zur Verfügung zu stellen, welche eine sichere und eindeutige Zuordnung der NMR-Signale zu den verschiedenen magnetisch aktiven Kernen ermöglicht und welche die Zahl der notwendigen manuellen Eingriffe verringert.

10

Diese Aufgabe der Erfindung wird durch ein Analysesystem zur automatisierten Analyse eines Satzes von NMR-Spektren gemäß Anspruch 1 sowie durch ein Verfahren zur automatisierten Analyse eines Satzes von NMR-Spektren gemäß Anspruch 18 gelöst.

15

Das erfindungsgemäße Analysesystem dient zur automatisierten Analyse eines Satzes von NMR-Spektren, welcher für eine n Aminosäuren umfassende Polypeptidkette aufgezeichnet wurde, und umfasst eine Bibliothek von Kreuzsignalmuster-Suchmasken, wobei eine Kreuzsignalmuster-Suchmaske zur spezifischen Erfassung der NMR-Signale eines Fragments der untersuchten Polypeptidkette vorgesehen ist. Es können dabei Kreuzsignale eines Fragments einer Aminosäure, oder von Fragmenten mehrerer sequentiell folgender, oder alle Kreuzsignale einer oder mehrerer Aminosäuren, die sequentiell in der Polypeptidkette verknüpft sind, erfasst werden. Die Fragmente können damit aus verbundenen Hauptkettenatomen, eventuell einschließlich der β -Kohlenstoffe, oder nur aus Seitenkettenatomen der einzelnen oder verknüpfter Aminosäuren bestehen. Des weiteren umfasst das Analysesystem Mittel zur Auswahl der zur Analyse benötigten Kreuzsignalmuster-Suchmasken der Bibliothek entsprechend der Primärsequenz der Polypeptidkette, welche für jedes in der Primärsequenz enthaltene Fragment die zugehörigen Kreuzsignalmuster-Suchmasken auswählen. Außerdem weist das Analysesystem Mittel zur Mustererkennung auf, welche die verschiedenen Ergebnisse der entsprechend der Primärsequenz der Polypeptidkette ausgewählten Kreuzsignalmuster-Suchmasken verbinden und mit dem Satz von NMR-Spektren korrelieren. Darüber hinaus umfaßt das Analysesystem Mittel zur Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der Polypeptidkette entsprechend dem Ergebnis der Mustererkennung.

20

25

30

Die erfindungsgemäße Lösung erlaubt insbesondere dann eine sehr zuverlässige Zuordnung, wenn für ein Protein ein Satz von NMR-Spektren aufgezeichnet wird, der neben 3D-Experimenten zur Zuordnung der Haupt- und Seitenkettensignale auch aminosäuretyp-

5 spezifische 2D-Experimente enthält. Diese aminosäuretyp-spezifischen NMR-Experimente enthalten nur Kreuzsignale einer oder mehrerer Aminosäuretypen, die entweder Korrelationen zwischen Signalen von Atomen in der Protein-Hauptkette oder von Seitenkettensignalen darstellen.

10 Die erfindungsgemäßen Kreuzsignalmuster-Suchmasken sind so gestaltet, dass Signalmuster von Fragmenten, welche zu bestimmten Aminosäuretypen oder Verknüpfungen davon gehören, durch eine kombinierte Anwendung auf aminosäuretyp-spezifische NMR-Experimente und 3D-Tripelresonanzexperimente erkannt werden können. Diese Erkennung der in den Fragmenten enthaltenen Signale bestimmter Aminosäuretypen wird insbesondere dann er-
15 folgreich sein, wenn diese zu Beginn der Suche in Form von zweidimensionalen Korrelationen von Signalen der Protein-Hauptkette, oder von Signalen der Seitenketten als Ausgangspunkt verwendet werden.

20 Mit den erfindungsgemäßen Kreuzsignalmuster-Suchmasken lassen sich damit insbesondere die Muster abfragen bzw. die Fragmente suchen, die von bestimmten Zweier- und Dreier-Kombinationen von Aminosäuren, wie z.B. dem Paar Valin-Threonin hervorgerufen werden. Solche Kombinationen kommen in der Polypeptidkette nur einmal oder wenige Male vor, so dass sich eine dementsprechend kurze Trefferliste ergibt. Diese hohe Selektivität einer für Zweier- bzw. Dreiergruppen von Aminosäuren spezifischen Kreuzsignalmuster-Suchmaske
25 ermöglicht in der überwiegenden Zahl der Fälle eine eindeutige Zuordnung zwischen den erfassten Kreuzsignalen und dem zugehörigen Fragment der Polypeptidkette. Vieldeutigkeiten bei der Zuordnung können so reduziert werden.

30 Nachdem mit Hilfe der Kreuzsignalmuster-Suchmasken Zuordnungen zwischen den Kreuzsignalen und den verschiedenen gefundenen Fragmenten andererseits vorgenommen wurden, müssen diese verschiedenen Teilzuordnungen in einem zweiten Schritt verbunden werden. Durch Verbinden der verschiedenen Teil-Zuordnungen gelangt man zu einer Zuordnung von allen Signalen des Spektrums zu den zugehörigen magnetisch aktiven Kernen der Polypeptid-

kette. Bei Verwendung der erfindungsgemäßen Kreuzsignalmuster-Suchmasken zur Erfassung von Fragmenten der untersuchten Polypeptidkette ergibt sich jeweils ein Überlapp der verschiedenen auszuwertenden Fragmente. Für die magnetisch aktiven Kerne im Überlappbereich müssen die jeweils mit verschiedenen Kreuzsignalmuster-Suchmasken ermittelten chemischen Verschiebungen übereinstimmen.

5 Mit Hilfe dieser Randbedingung lassen sich die verschiedenen Teil-Zuordnungen verbinden, wobei insbesondere auf die chemischen Verschiebungen von H_N sowie N abgestellt wird. Wegen des Überlapps zwischen den Kreuzsignalmuster-Suchmasken lässt sich mit den erfindungsgemäßen Kreuzsignalmuster-Suchmasken die Verbindung der aus selektiven Suchen erhaltenen Teil-Zuordnungen zu einer 10 Gesamt-Zuordnung wesentlich einfacher und sicherer bewerkstelligen, als dies im Stand der Technik möglich war. Die erfindungsgemäßen Kreuzsignalmuster-Suchmasken bieten insofern sowohl bei der eigentlichen Peakzuordnung als auch bei der nachfolgenden Verbindung der Teil-Zuordnungen Vorteile gegenüber den einfachen Kreuzsignalmuster-Suchmasken des 15 Stands der Technik.

15 Insgesamt wird durch den Einsatz der erfindungsgemäßen Kreuzsignalmuster-Suchmasken, mit denen sich die NMR-Signale eines Fragments der untersuchten Polypeptidkette spezifisch erfassen lassen, eine höhere Sicherheit bei der Zuordnung sowie eine bessere Erkennungsrate ermöglicht. Da die Zahl der bei der Zuordnung auftretenden Vieldeutigkeiten verringert wird, 20 muss im Verlauf der Zuordnung weniger häufig manuell eingegriffen werden. Insofern stellt die Erfindung einen wichtigen Schritt beim Übergang von der semiautomatischen zur vollautomatischen Zuordnung von NMR-Signalpeaks dar. Sobald eine verlässliche, vollautomatische Zuordnung möglich ist, lässt sich der Durchsatz bei der Bestimmung von Proteinstrukturen signifikant erhöhen. Außerdem steigt die Verlässlichkeit der so erhaltenen Strukturdaten 25 an.

30 Mit Hilfe der erfindungsgemäßen Kreuzsignalmuster-Suchmasken wird darüber hinaus eine geschickte und instruktive Kapselung der Analysetools für die Zuordnung der Kreuzsignalmuster erreicht. Dadurch können die Ergebnisse der Zuordnung auch besser nachvollzogen werden.

Es ist von Vorteil, wenn die Fragmente der untersuchten Polypeptidkette jeweils zwei bzw. drei bestimmte aufeinanderfolgende Aminosäuren umfassen. Mit einer erfindungsgemäßen

Kreuzsignalmuster-Suchmaske kann ein bestimmtes Fragment von zwei (oder drei) Aminosäuren eindeutig anhand seiner NMR-Signale identifiziert werden. Mit den erfindungsgemäßen Kreuzsignalmuster-Suchmasken lassen sich damit insbesondere die Muster abfragen bzw. die Fragmente suchen, die von bestimmten Zweier- und Dreier-Kombinationen von Aminosäuren, wie z.B. dem Paar Valin-Threonin hervorgerufen werden. Solche Kombinationen kommen in der Polypeptidkette nur einmal oder wenige Male vor, so dass sich eine dementsprechend kurze Trefferliste ergibt. Diese hohe Selektivität einer für Zweier- bzw. Dreiergruppen von Aminosäuren spezifischen Kreuzsignalmuster-Suchmaske ermöglicht in der überwiegenden Zahl der Fälle eine eindeutige Zuordnung zwischen den erfassten Kreuzsignalen und dem zugehörigen Fragment der Polypeptidkette. Vieldeutigkeiten bei der Zuordnung können so reduziert werden.

Es ist von Vorteil, wenn der Satz von NMR-Spektren sowohl NMR-Experimente zur Analyse der Hauptkettensignale als auch NMR-Experimente zur Analyse der Seitenkettensignale umfasst. Die Verknüpfung zwischen den Seitenketten und der Hauptkette wird insbesondere über die chemische Verschiebung der C_{α} - sowie der C_{β} -Kerne hergestellt. Mit Hilfe der erfindungsgemäßen Kreuzsignalmuster-Suchmasken, welche zur spezifischen Erfassung der NMR-Signale eines Fragments der untersuchten Polypeptidkette vorgesehen sind, können die Haupt- und Seitenkettensignale gemeinsam ausgewertet werden.

Es ist von Vorteil, wenn die zur Analyse der Hauptkettensignale verwendeten NMR-Experimente 3D-Experimente, und insbesondere 3D-Experimente der Typen CBCA(CO)NNH, CBCANNH, HA(CO)NNH, HANNH, HAHB(CO)NNH, HAHBNNH, HN(CA)CO, HNCO, HN(CO)CA, HNCA umfassen. Bei den angeführten Experimenten handelt es sich um 3D-Experimente, mit denen sich die chemischen Verschiebungen der magnetisch aktiven Kerne der Hauptkette erfassen lassen. Die Vielfalt der verfügbaren 3D-Experimente erlaubt aber auch eine mehrfache Absicherung der Ergebnisse.

Es ist von Vorteil, wenn die zur Analyse der Seitenkettensignale verwendeten NMR-Experimente Experimente der Typen HCCH-COSY, HCCH-TOCSY, HCC(CO)NH-TOCSY umfassen.

Es ist von Vorteil, wenn die zur Analyse der Haupt- und Seitenkettensignale verwendeten NMR-Experimente aminosäuretyp-spezifische $^1\text{H}/^{15}\text{N}$ -Experimente umfassen, welche selektiv sind für einen Aminosäuretyp oder für eine Gruppe von Aminosäuretypen. Für ein Protein wird ein Satz von NMR-Spektren aufgezeichnet, der neben den 3D-Experimenten zur Erfassung der Hauptkettenstruktur auch aminosäuretyp-spezifische 2D-Experimente enthält.

5 Diese aminosäuretyp-spezifischen NMR-Experimente enthalten nur Kreuzsignale einer oder mehrerer Aminosäuretypen, die entweder Korrelationen zwischen Signalen von Atomen in der Protein Hauptkette oder von Seitenkettensignalen darstellen. Aminosäuretyp-spezifische 2D-Experimente ermöglichen die selektive Anregung der Seitenketten eines Aminosäuretyps

10 bzw. einer Gruppe von Aminosäuretypen. Anschließend wird die Magnetisierung über die Seitenkette zu den Hauptketten-Stickstoffen und Amidprotonen transferiert. Mit Hilfe von aminosäuretyp-spezifischen $^1\text{H}/^{15}\text{N}$ -Experimenten lassen sich die von einem bestimmten Aminosäuretyp bzw. einer Gruppe von Aminosäuretypen verursachten NMR-Signale, die eine Art „Fingerabdruck“ eines bestimmten Aminosäuretyps bzw. einer Gruppe von Amino-

15 säuren darstellen, hochspezifisch erfassen. Mit einer erfindungsgemäßen Kreuzsignal-muster-Suchmaske können dann gezielt die NMR-Signale eines Fragments der untersuchten Polypeptidkette abgefragt werden. Dadurch wird eine höhere Sicherheit bei der Zuordnung sowie eine bessere Erkennungsrate ermöglicht.

20 Die zur Durchführung der aminosäuretyp-spezifischen 2D-Experimente benötigten Pulssequenzen lassen sich auf einfache Weise aus den vor allem zur Bestimmung der Hauptkettenstruktur verwendeten Tripelresonanz-Experimenten ableiten.

Es ist von Vorteil, wenn die zur Analyse der Haupt- und Seitenkettensignale erforderlichen 25 aminosäuretyp-spezifischen 2D-Experimente entsprechend der Primärsequenz der Polypeptidkette festgelegt werden. Diese Primärsequenz ist vorher bekannt. Die zur Durchführung der NMR-Messungen erforderliche Menge an Protein wird nämlich unter Zuhilfenahme einer entsprechenden DNS-Sequenz mit Hilfe von biotechnologischen Methoden erzeugt. Wenn nun beispielsweise in der Primärsequenz der Polypeptidkette die Aminosäure Cystein nicht 30 vorkommt, so ist es auch nicht notwendig, das aminosäuretyp-spezifische 2D-Experiment für Cystein durchzuführen. Anhand der Primärsequenz lässt sich also der minimal erforderliche Datensatz für die NMR-Experimente festlegen.

Es ist von Vorteil, wenn die zur Analyse der Haupt- und Seitenketten signale verwendeten NMR-Experimente eine Kombination von 2D- und 3D-Experimenten umfassen. Die kombinierte Verwendung von Hauptkettenexperimenten und aminosäuretyp-spezifischen 2D-Experimenten, zusammen mit die Auswertung der NMR-Signale mit Hilfe von Kreuzsignal-
5 muster-Suchmasken, ermöglicht eine erhebliche Performancesteigerung bei der automatisier-ten Auswertung von NMR-Spektren. Die erfundungsgemäßen Kreuzsignalmuster-Suchmasken sind so gestaltet, dass Signalmuster von Fragmenten, welche zu bestimmten Aminosäuretypen oder Verknüpfungen davon gehören, durch eine kombinierte Anwendung auf aminosäuretyp-spezifische NMR-Experimente und 3D-Tripelresonanzexperimente er-
10 kannt werden können. Da insbesondere mit Hilfe von aminosäuretyp-spezifischen 2D-Experimenten die Zahl der bei der Zuordnung auftretenden Vieldeutigkeiten verringert wer-
den kann, muss im Verlauf der Zuordnung weniger häufig manuell eingegriffen werden. Die Erfundung stellt daher einen wichtigen Schritt beim Übergang von der semiautomatischen zur vollautomatischen Auswertung von NMR-Spektren dar.

15 Gemäß einer vorteilhaften Ausführungsform der Erfundung werden ausgehend von der Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der Polypeptidkette die chemischen Verschiebungen zusammengefasst und auf ihre Richtigkeit überprüft. Ausgehend von der Zuordnung der NMR-Signale können beispielsweise die für die verschiedenen ma-
20 gnetisch aktiven Kerne einer Aminosäure ermittelten chemischen Verschiebungen in Form eines Vektors zusammengefasst werden. Hinsichtlich der Hauptkette kann dann insbesondere mittels der chemischen Verschiebung von H_N und N, welche in verschiedenen Experimenten unabhängig voneinander ermittelt werden, eine Konsistenzüberprüfung durchgeführt werden. Für die in verschiedenen Experimenten ermittelten chemischen Verschiebungen müssen sich
25 übereinstimmende bzw. nahe beieinander liegende Werte ergeben. Entsprechend erlauben die chemischen Verschiebungen der Kerne C_α und C_β , welche sowohl in Hauptkettenexperi-
menten als auch in Seitenkettenexperimenten ermittelt werden, eine derartige Konsistenzüberprü-
fung für die Haupt- und Seitenketten.

30 Gemäß einer weiteren Ausführungsform der Erfundung umfasst der Satz von NMR-Spektren Spektren des Typs NOESY, deren Auswertung insbesondere Informationen über die Ab-
stände der verschiedenen Kerne der Polypeptidkette voneinander liefert. In Experimenten des Typs NOESY wird die durch den Kern-Overhauser-Effekt verursachte Kreuzrelaxation

erfasst. Aus den Amplituden der NOE-Kreuzsignale lassen sich Rückschlüsse auf die Abstände zwischen den beteiligten Kernen ziehen. Für die Ermittlung der Proteinstruktur sind Spektren des Typs NOESY daher von besonderer Bedeutung.

5 Gemäß einer weiteren Ausführungsform der Erfindung erfolgt die Zuordnung der NMR-Spektren des Typs NOESY zu den verschiedenen Kernen der Polypeptidkette anhand der für die Kerne ermittelten chemischen Verschiebungen. Die Zuordnung der Kreuzsignale in den NOESY-Spektren zu den verschiedenen magnetisch aktiven Kernen wird insbesondere anhand der protonenchemischen Verschiebungen vorgenommen. Aber auch wenn die protonenchemischen Verschiebungen mit zufriedenstellender Genauigkeit vorab bestimmt wurden und daher vorbekannt sind, bleiben wegen der Mehrfachbedeutung der einzelnen Kreuzsignale noch Vieldeutigkeiten bestehen. Aus diesem Grund ist es um so wichtiger, anhand von möglichst genau erfassten protonenchemischen Verschiebungen einen möglichst großen Anteil der NOESY-Spektren eindeutig zuordnen zu können.

10

15 Gemäß einer weiteren vorteilhaften Ausführungsform der Erfindung dienen die bei der Auswertung der NMR-Spektren erhaltenen Größen als Eingangsgrößen für Strukturberechnungs-Software. Wichtige Eingangsgrößen für Strukturberechnungs-Programme sind vor allem die aus den zugeordneten NOESY-Spektren erhaltenen Kernabstände. Dem Strukturberechnungs-Programm kann hierzu eine Liste der Amplituden der NOE-Kreuzsignale und der Frequenzkoordinaten der Peaks sowie die Resonanzzuordnung übergeben werden. Weitere Eingangsgrößen können die Kopplungskonstanten zwischen verschiedenen Kernen sein, denn aus diesen Kopplungskonstanten lassen sich die Diederwinkel zwischen verschiedenen Kernen ermitteln.

20

25 Es ist von Vorteil, wenn die Kreuzsignalmuster-Suchmasken jeweils eine Anzahl von vordefinierten Signalsuchbereichen umfassen, wobei durch das Auftreten von NMR-Signalen innerhalb der Bereichsgrenzen eines Signalsuchbereichs die Wahrscheinlichkeit erhöht wird, dass das durch die Kreuzsignalmuster-Suchmaske definierte Signalmuster vorliegt. Auf diese Weise lässt sich das zu suchende Peakmuster exakt mittels einer Anzahl von Suchbereichen definieren. Dadurch wird eine echte Mustererkennung möglich. Jeder innerhalb der Grenzen eines Suchbereichs auftretende Signalpeak erhöht die Bewertungszahl für die Kreuzsignalmuster-Suchmaske. Dabei ist es insbesondere von Vorteil, dass auch bei Fehlen einzelner Peaks des

Signalmusters ein Signalmuster noch erkannt werden kann, wenn es ansonsten hinreichend gut mit der Kreuzsignalmuster-Suchmaske übereinstimmt. Bei der Bewertung, ob ein Peakmuster mit der Kreuzsignalmuster-Suchmaske übereinstimmt, kommt es auf die insgesamt festgestellten Übereinstimmungen an.

5

Dabei ist es von Vorteil, wenn die Kreuzsignalmuster-Suchmasken jeweils eine Anzahl von vordefinierten Leerbereichen umfassen, wobei durch das Fehlen von NMR-Signalen innerhalb der Bereichsgrenzen eines Leerbereichs die Wahrscheinlichkeit erhöht wird, dass das durch die Kreuzsignalmuster-Suchmaske definierte Signalmuster vorliegt. Die Definition von

10 Leerbereichen ist beispielsweise dann sinnvoll, wenn zwei verschiedene Seitenkettenstrukturen zu zwei einander ähnlichen Signalmustern führen, wobei das zweite Signalmuster einige zusätzliche Peaks aufweist, welche im ersten Signalmuster nicht enthalten sind. Das Fehlen dieser Signalpeaks ist dann genau das typische an dem ersten Signalmuster, so dass es sich zur Erfassung des ersten Signalmusters empfiehlt, an den entsprechenden Stellen Leerbereiche zu

15 definieren. Wenn dann innerhalb der Grenzen der Leerbereiche keine Peaks auftreten, so wird die Bewertungszahl für das Vorliegen des ersten Signalmusters erhöht. Durch die Definition von Leerbereichen können die beiden Signalmuster deshalb besser unterschieden werden.

20 Es ist von Vorteil, wenn ausgehend von der erwarteten Anzahl der NMR-Signale in den Spektren durch Iteration die Schwellwerte und Suchbereiche für die Kreuzsignalmuster-Suchmasken ermittelt werden. Bei dieser Vorgehensweise werden die Suchbereiche anfangs durch weit gesteckte Grenzen definiert, während die Schwellwerte relativ niedrig gewählt werden. Die Auswertung der beim ersten Durchlauf aufgefundenen Signalpeaks liefert erste

25 Aussagen darüber, welches bzw. welche Kreuzsignalmuster voraussichtlich vorliegen. In einer zweiten, modifizierten Suche kann dann speziell nach diesen wahrscheinlichsten Kandidaten gesucht werden, dabei können die Suchbereiche in Abhängigkeit von den chemischen Verschiebungen der bei der ersten Suche aufgefundenen Peaks verkleinert bzw. verschoben werden, um die Suche zu verfeinern. Außerdem ist es möglich, bei der zweiten Suche mit erhöhten Schwellwerten zu arbeiten. Mit Hilfe dieser iterativen Vorgehensweise lässt sich erreichen,

30 dass die Kreuzsignalmuster-Suche schrittweise in Richtung des tatsächlich vorliegenden Kreuzsignalmusters konvergiert.

Gemäß einer weiteren vorteilhaften Ausführungsform der Erfindung umfasst die Kreuzsignalmuster-Suchmaske eine Vielzahl von Unter-Suchmasken zur Analyse der verschiedenen NMR-Spektren des aufgezeichneten Satzes von NMR-Spektren. Die eigentliche Kreuzsignalmuster-Suchmaske ergibt sich insofern als Gesamtheit verschiedener Unter-Suchmasken,

5 welche jeweils verschiedene zwei-, drei- oder höherdimensionale Spektren absuchen. Der Satz von NMR-Spektren wird also mit einem entsprechenden Satz von Unter-Suchmasken analysiert. Dies hat insbesondere den Vorteil, dass sich die Modifikation von Suchbereichsgrenzen simultan auf alle Unter-Suchmasken auswirkt. Dadurch wird die Handhabung des Satzes von Suchmasken vereinfacht.

10

Bei dem erfindungsgemäßen Verfahren zur automatisierten Analyse eines Satzes von NMR-Spektren, welcher für eine n Aminosäuren umfassende Polypeptidkette aufgezeichnet wurde, werden zunächst die zur Analyse benötigten Kreuzsignalmuster-Suchmasken aus einer Bibliothek von Kreuzsignalmuster-Suchmasken ausgewählt, wobei eine Kreuzsignalmuster-

15 Suchmaske zur spezifischen Erfassung der NMR-Signale eines Fragments der untersuchten Polypeptidkette vorgesehen ist, und wobei das Auswählen der benötigten Kreuzsignalmuster-Suchmasken entsprechend den in der Primärsequenz enthaltenen Fragmenten erfolgt. Anschließend wird eine Mustererkennung durchgeführt, indem die verschiedenen ausgewählten Kreuzsignalmuster-Suchmasken mit dem Satz von NMR-Spektren korreliert werden. Entsprechend dem Ergebnis dieser Mustererkennung werden die NMR-Signale zu den verschiedenen Spinsystemen der Polypeptidkette zugeordnet.

20

25 Da die erfindungsgemäßen Kreuzsignalmuster-Suchmasken jeweils sämtliche NMR-Signale eines Fragments der untersuchten Polypeptidkette zusammen auswerten, können mit dem erfindungsgemäßen Verfahren die NMR-Signale eines Fragments, insbesondere eines zwei bzw. drei Aminosäuren umfassenden Fragments, hochselektiv erfasst werden. Wenn dann für die einzelnen Fragmente der Polypeptidkette verlässliche Signalzuordnungen vorliegen, kann wegen des Überlapps zwischen den Analyseergebnissen eine Gesamtzuordnung der NMR-Signale der Polypeptidkette hergeleitet werden. Gegenüber den Verfahren des Stands der

30 Technik wird die Eindeutigkeit und Sicherheit der Zuordnungen verbessert. Manuelle Eingriffe des Benutzers sind weniger häufig erforderlich und deshalb eignet sich das Verfahren insbesondere zur vollautomatischen Auswertung von NMR-Spektren. Dadurch kann der

Zeit- und Geldbedarf bei der Ermittlung von Proteinstrukturen durch NMR-Spektroskopie weiter abgesenkt werden.

Nachfolgend wird die Erfindung anhand eines in der Zeichnung dargestellten Ausführungs-

5 beispiels weiter beschrieben. Es zeigen:

Fig. 1 einen Ablaufplan für die Aufzeichnung der benötigten NMR-Spektren;

10 Fig. 2 eine Tabelle, aus der die in der Aminosäuresequenz enthaltenen Paare von Aminosäuren ausgelesen werden können;

15 Fig. 3 eine Reihe von Beispielen für aminosäuretyp-spezifische 2D-Experimente, mit denen gezielt das Vorliegen oder Nichtvorliegen bestimmter Seitenkettenstrukturen abgefragt werden kann;

Fig. 4 eine Kreuzsignalmuster-Suchmaske, mit der die in den Spektren enthaltenen Signalmuster erkannt und ausgewertet werden können; sowie

Fig. 5A,

20 Fig. 5B einen Ablaufplan für die Auswertung der aufgezeichneten NMR-Spektren, wobei eine Zuordnung zwischen den auftretenden Signalpeaks und den Spinsystemen des Proteins hergestellt wird.

In Fig. 1 ist der Ablauf bei der Aufzeichnung der benötigten NMR-Spektren dargestellt. Ausgangspunkt für die Festlegung der erforderlichen Experimente ist die Primärstruktur des Proteins. Meist werden die benötigten Proteine mit Hilfe von biotechnologischen Methoden unter Zuhilfenahme von entsprechenden DNS-Abschnitten synthetisiert, da sich auf diese Weise die benötigten Proteinmengen einfach herstellen lassen. Es ist daher im folgenden davon auszugehen, dass die Primärstruktur des Proteins vorbekannt ist, und dass die NMR-spektroskopischen Experimente lediglich zur Strukturmittlung des Proteins eingesetzt werden sollen.

Im Schritt 1 werden die in der Primärsequenz enthaltenen Paare von aufeinanderfolgenden Aminosäuren ermittelt und aufgelistet. Da die Primärstruktur bekannt ist, kann dies auf sehr einfache Weise von einem selbstgeschriebenen Computerprogramm namens „selma“ durchgeführt werden. In Fig. 2 ist das von dem Programm „selma“ für das Protein OPR gelieferte

5 Ergebnis dargestellt. Die entlang der x- und der y-Achse aufgetragenen Buchstaben bezeichnen jeweils die 20 möglichen Aminosäuren. Die entlang der y-Achse aufgelisteten Aminosäuren bezeichnen die in der ersten Position des betrachteten Aminosäure-Paars vorkommende Aminosäure, während die entlang der x-Achse aufgelisteten Aminosäuren die Aminosäure an der zweiten Stelle des Aminosäure-Paars bezeichnen.

10

In der sich ergebenden Matrix sind die verschiedenen in der Aminosäuresequenz des Proteins OPR auftretenden Aminosäure-Paare eingetragen. Der Tabelle kann entnommen werden, dass das Aminosäure-Paar AE genau einmal in der Sequenz vorkommt, während das Aminosäure-Paar EA nicht in der Sequenz enthalten ist. Falls die Ziffer 2 oder 3 an einer bestimmten Matrixposition eingetragen ist, so bedeutet dies, dass das entsprechende Aminosäure-Paar in der Sequenz mehrfach auftritt. Dies ist z.B. beim Aminosäure-Paar ED der Fall.

15

Die so erhaltenen Informationen über die in der Primärsequenz enthaltenen Aminosäure-Paare dienen im Schritt 2 zur Festlegung eines Satzes von NMR-Experimenten, welche zur 20 Ermittlung der Proteinstruktur durchgeführt werden müssen. Das Ziel ist, möglichst keine überflüssigen Experimente durchzuführen, welche nur unnötig die Messzeit verlängern würden. So lässt sich beispielsweise anhand der in Fig. 2 gezeigten Ergebnisse des Programms „selma“ erkennen, dass die Aminosäure Cystein im Protein OPR nicht vorkommt. Die Aufzeichnung eines Cystein-selektiven 2D-Experiments ist nicht notwendig, und das Cystein-selektive 2D-Experiment wird daher auch nicht Bestandteil des in Schritt 2 festgelegten Satzes von NMR-Experimenten.

25

In Schritt 3 werden die zwei- bzw mehrdimensionalen NMR-Spektren des Satzes in einem NMR-Spektrometer aufgezeichnet. Das Spektrometer umfasst eine Spektrometersteuerung, welche anfangs die Betriebsparameter des Spektrometers, wie beispielsweise die Protonenträgerfrequenz sowie die Länge von 90°-Pulsen ermittelt und einstellt. Anschließend werden die in Schritt 2 festgelegten Spektren der Reihe nach aufgezeichnet. Hierzu enthält die Spektrometersteuerung eine Auswahl an standardisierten NMR-Pulssequenzen.

Die Aufzeichnung der für ein Protein benötigten NMR-Spektren erfordert eine Messzeit von einigen Wochen. Danach stehen die aufgezeichneten Spektren als Dateien eines Projektordners 4 zur Verfügung und können im Schritt 5 weiter ausgewertet werden. Als Ergebnis erhält

5 man 2D-Spektren „2rr“ sowie 3D-Spektren „3rrr“. Aus den im Projektordner 4 enthaltenen Dateien ermittelt das Programm „gnat“ verschiedene für die weitere Auswertung benötigte statistische Parameter 6, die beispielsweise für die Festlegung von Schwellwerten benötigt werden. Mit Hilfe derartiger Schwellwerte können die in den Spektren enthaltenen Peaks vom Untergrundrauschen unterschieden werden.

10

Zur Erfassung der chemischen Verschiebungen der magnetisch aktiven Kerne der Hauptkette werden typischerweise 3D-Tripelresonanzexperimente sowie aminosäuretyp-spezifische 2D-Experimente verwendet, mit denen sich insbesondere auch die Resonanzen von ^{13}C und ^{15}N -Kernen erfassen lassen. Hierzu dienen insbesondere folgende Paare von 3D-Experimenten:

15

CBCA(CO)NNH und CBCANNH, HA(CO)NNH und HANNH, HAHB(CO)NNH und HAHBNNH, HN(CA)CO und HNCO sowie HN(CO)CA und HNCA. In Klammern angegebene Kerne (z.B. 3D-HN(CO)CA) werden nicht detektiert, sind aber bei einem Kohärenztransfer involviert.

20

Zur Ermittlung der benötigten chemischen Verschiebungen würde bereits die Durchführung einer geringen Zahl von Tripelresonanzexperimenten ausreichen. Um verlässliche Ergebnisse zu erzielen, ist es allerdings von Vorteil, die verschiedenen im Protein auftretenden Korrelationen jeweils durch mehrere Experimente zu erfassen, so dass die Ergebnisse auf ihre Konsistenz überprüft werden können.

25

Mit Hilfe des Experiments CACBNNH erhält man die Frequenzen der Kerne H_N , N , C_α und C_β der i-ten Aminosäure sowie die Frequenzen der Kerne C_α und C_β der Aminosäure i-1. Das CACB(CO)NH-Spektrum enthält die Korrelationen der Kerne H_N und N der Aminosäure i mit den Kernen C_α und C_β der Aminosäure i-1.

30

Mit dem 3D-Experiment des Typs HNCA werden die Frequenzen der Kerne N , H_N und C_α der i-ten Aminosäure erfasst. Entsprechend können mittels des Experiments HNCA auch die

Frequenzen der Kerne N , H_N sowie C_α der Aminosäure $i+1$ erfasst werden. Mit dem Experiment HNCA ist es darüber hinaus auch möglich, den Kern C_α der Aminosäure i mit den Kernen N und H_N der Aminosäure $i+1$ zu korrelieren.

5 Der wesentliche Vorteil der Tripelresonanzexperimente ist die verhältnismäßig geringe Anzahl von Signalpeaks in den Spektren. Tripelresonanzexperimente enthalten pro Aminosäure nur ein oder maximal zwei Kreuzsignale. Dadurch wird eine weitgehende Automatisierung der Auswertung möglich.

10 Die frequenzspezifische Zuordnung erfolgt somit durch die Kombination einzelner „Bausteine“, welche sich in Teilen überschneiden. Besonders wichtig sind solche Experimente, die eine Erfassung der chemischen Verschiebungen der β -Kohlenstoffe erlauben.

Als besonders vorteilhaft für die Aufklärung der Resonanzzuordnung hat sich die Verwendung der Pulssequenz MUSIC (Multiplicity Selective In-Phase Coherence Transfer) erwiesen. 15 MUSIC-Pulssequenzen zur gezielten Anregung der Seitenketten einer Gruppe von Aminosäuren bzw. eines bestimmten Aminosäuretyps können durch Abwandlung der Pulssequenzen von 3D-Tripelresonanzexperimenten (wie beispielsweise CBCA(CO)NNH) erhalten werden.

20 In Fig. 3 ist der jeweilige Magnetisierungstransfer für eine Reihe von aminosäuretyp-
spezifischen 2D-Experimenten dargestellt. Zunächst wird eine bestimmte in der Seitenkette befindliche Gruppe durch die MUSIC-Sequenz angeregt. Bei den gezeigten Beispielen wird jeweils die durch eine rechteckige Umrandung hervorgehobene CH_2 - bzw. CH_3 -Gruppe ange-
rengt. Von dort wird die Magnetisierung entlang der Seitenkette zu dem C_α -Atom und von dort weiter zum Amidproton N transferiert. Der Unterschied zwischen den in der linken und in der rechten Spalte gezeigten Experimenten besteht in der Art des Transfers von dem C_α -Atom zu dem Stickstoff N . Bei den in der linken Spalte gezeigten Experimenten gelangt die Magnetisierung von dem C_α -Kern zu der Carbonylgruppe und von dort weiter zum Stickstoff 25 N und zum Amidproton H_N der Aminosäure $i+1$. Wegen dieses Magnetisierungstransfers zur nächstfolgenden Aminosäure werden diese Experimente als $(i+1)$ -HSQCs bezeichnet. Bei den in der rechten Spalte gezeigten $(i, i+1)$ -HSQCs dagegen gelangt die Magnetisierung von 30

dem C_{α} -Kern aus entweder zum Stickstoff N derselben Aminosäure i oder zum Stickstoff N der benachbarten Aminosäure i+1. Mit Hilfe der in Fig. 3 gezeigten Experimente lassen sich 2D-Spektren selektiv für einen bestimmten Typ von Aminosäure (z. B. für Ser, 1. Zeile; Leu, 3. Zeile) oder selektiv für bestimmte Gruppen von Aminosäuren (vgl. Ile/Val, 2. Zeile; 5 Asp/Asn, 4. Zeile sowie Glu/Gln, 5. Zeile) erfassen.

Die aufgezeichneten 2D- sowie 3D-Spektren werden anschließend einer Mustererkennung unterzogen, um die in den Spektren auftretenden Signalpeaks den einzelnen Spinsystemen des Proteins zuordnen zu können. Bei den Lösungen des Stands der Technik wurde diese Zuordnung 10 meist mit Hilfe von Peaksignalisten vorgenommen. Gegenüber derartigen Lösungsansätzen bietet die Verwendung von Mustererkennungs-Routinen insofern Vorteile, als hier die gemessenen Kreuzsignalmuster in ihrer Gesamtheit beurteilt werden können.

In Fig. 4 ist eine zur Analyse von Kreuzsignalmustern verwendete Kreuzsignalmuster-Suchmaske 15 dargestellt. An den Positionen 8, 10, 12 sowie den hierzu spiegelbildlichen Positionen 14, 16 und 18 des zweidimensionalen Spektrums werden Signalpeaks erwartet. Um die erwarteten Peakpositionen herum sind rechteckige Signalsuchbereiche 9, 11, 13 sowie 15, 17, 19 definiert. Innerhalb der so definierten Bereiche auftretende Signalpeaks werden von der Mustererkennungs-Software erfasst, während außerhalb der Signalsuchbereiche auftretende 20 Peaks nicht erfasst werden. Die Kreuzsignalmuster-Suchmaske umfasst also die vordefinierten Signalsuchbereichen 9, 11, 13, 15, 17, 19, wobei die Software innerhalb dieser Signalsuchbereiche nach den erwarteten Signalpeaks sucht.

Der Ablauf der Mustererkennung und Zuordnung ist in den Fig. 5A und 5B dargestellt. Zur 25 Analyse der Kreuzsignalmuster dient eine Bibliothek 19 von erfindungsgemäßen Kreuzsignalmuster-Suchmasken, welche der Mustererkennungssoftware 20 zur Verfügung steht. Mit jeder der erfindungsgemäßen Kreuzsignalmuster-Suchmasken lassen sich die Signalpeaks eines Fragments von zwei (oder drei) aufeinanderfolgenden Aminosäuren erfassen und zuordnen. Die Auswahl der für die Mustererkennung benötigten Kreuzsignalmuster-Suchmasken 30 erfolgt dabei entsprechend der durch das Programm „selma“ vorgenommenen Zerlegung der Aminosäuresequenz in Zweierfragmente. Die zur Zuordnung der Signalpeaks des Proteins benötigten Kreuzsignalmuster-Suchmasken werden entsprechend den in der Primärsequenz

enthalteten Zweiergruppen aus der Bibliothek 19 ausgewählt und der Mustererkennungs-Software 20 zur Verfügung gestellt.

Mit Hilfe einer vordefinierten Kreuzsignalmuster-Suchmaske lassen sich sämtliche Signalpeaks, die von einer Zweiergruppe von aufeinanderfolgend angeordneten Aminosäuren verursacht werden, also sowohl die von den beiden Seitenketten verursachten Signalpeaks als auch die vom Hauptkettenfragment verursachten Signalpeaks, gemeinsam auswerten.

Es soll im folgenden davon ausgegangen werden, dass zur Zuordnung der Hauptkettensignale drei verschiedene Paare von Tripelresonanzexperimente, nämlich die Paare CBCANNH/CBCA(CO)NNH, HNCO/HN(CA)CO sowie HANNH/HA(CO)NNH, und die aminosäuretyp-spezifischen 2D-Experimente herangezogen werden. In diesem Fall enthält die Bibliothek 19 von Kreuzsignalmuster-Suchmasken insgesamt $3 \times 20 \times 20 = 3 \times 400 = 1200$ verschiedene Kreuzsignalmuster-Suchmasken, nämlich

- 400 Kreuzsignalmuster-Suchmasken zur Analyse von:
CBCANNH + CBCA(CO)NNH + zwei aminosäuretyp-spezifische 2D-Experimente,
- 400 Kreuzsignalmuster-Suchmasken zur Analyse von:
HNCO + HN(CA)CO + zwei aminosäuretyp-spezifische 2D-Experimente, sowie
- 400 Kreuzsignalmuster-Suchmasken zur Analyse von:
,HANNH + HA(CO)NNH + zwei aminosäuretyp-spezifische 2D-Experimente.

Jede der erfindungsgemäßen Kreuzsignalmuster-Suchmasken dient zur Auswertung von zwei aminosäuretyp-spezifischen 2D-Experimenten sowie von einem Paar von 3D-Tripelresonanzexperimenten. Um die verschiedenen Spektren mit einer Kreuzsignalmuster-Suchmaske auswerten zu können, umfasst die Kreuzsignalmuster-Suchmaske einen Satz von Unter-Suchmasken, wobei mit jeder Unter-Suchmaske ein bestimmter Typ von Spektren ausgewertet werden kann. Programmtechnisch präsentiert sich die Kreuzsignalmuster-Suchmaske jedoch als eine Einheit. Es ist daher möglich für die gesamte Kreuzsignalmuster-Suchmaske mit all ihren Unter-Suchmasken einen bestimmten Auswertungsparameter, beispielsweise eine Suchbereichsgrenze, zu verändern. Die Veränderung wirkt sich dann in selbstkonsistenter Art und Weise auf die Suchbereichsgrenzen in sämtlichen Unter-Suchmasken aus.

Die in der Bibliothek 19 enthaltenen, vordefinierten Kreuzsignalmuster-Suchmasken legen den Suchalgorithmus für das Auffinden bestimmter Signalmuster fest. Allerdings sind die Kreuzsignalmuster-Suchmasken in parameterunabhängiger Form niedergelegt. Die benötigten Suchbereichsgrenzen 22 sowie die zur Unterscheidung der Signalpeaks vom Rauschuntergrund benötigten Schwellwerte 23 werden den Kreuzsignalmuster-Suchmasken extern zur Verfügung gestellt.

Diese Vorgehensweise bietet die Möglichkeit, die bei der Durchführung einer Kreuzsignalmustersuche verwendeten Parameter und Suchbereiche im Laufe der Suche zu verändern und an neu gewonnene Informationen anzupassen. Insbesondere ist es von Vorteil, die Suchbereichsgrenzen anfangs sehr breit zu wählen und in Abhängigkeit von den innerhalb der Suchbereiche auftretenden Peaks iterativ zu verkleinern, um so das gesuchte Kreuzsignalmuster mit höherer Sicherheit erfassen zu können. Entsprechend können die Schwellwerte 23 von einem anfangs niedrigen Wert im Laufe der Suche auf immer höhere Werte angehoben werden, um so die Kreuzsignalmuster mit den höchsten Bewertungszahlen herauszufiltern.

Die im folgenden aufgeführten Programmroutine stellt eine Implementierung einer Kreuzsignalmuster-Suchmaske dar, welche die zwei seitenkettenspezifischen 2D-Experimente „sHSQCcoN“ sowie „sHSQCcaS“ sowie das Paar von 3D-Tripelresonanzexperimenten „HNcoCACB“ sowie „HNCACB“ auswertet. Diese für ein Paar von Aminosäuren spezifische Kreuzsignalmuster-Suchmaske umfasst also vier Unter-Suchmasken zur Auswertung der verschiedenen 2D- sowie 3D-Spektren. Die Verbindung der durch Auswertung der verschiedenen Spektren gefundenen Ergebnisse erfolgt durch die chemischen Verschiebungen der Kerne im Überlappbereich, also insbesondere über die chemischen Verschiebungen von H_N , N , C_α und C_β .

(----- Appendix 1 -----)

Aus dem Programmcode ist insbesondere erkennbar, dass die Kreuzsignalmuster-Suchmaske in ihrer abstrakt festgelegten Form noch keine numerisch festgelegten Suchbereiche und Schwellwerte aufweist. Die entsprechenden Variablen „submatrix_sizes“, „sweep_widths“, „ppm_offsets“ sowie „nucleus_species“ sind lediglich in abstrakter Form definiert.

Das folgende Programmlisting zeigt die Kreuzsignalmuster-Suchmaske für die 2D-Spektren „sHSQCcoN“, „sHSQCcaS“ sowie der 3D-Spektren „HNcoCACB“, „HNCACB“, wobei inzwischen die Suchbereichsgrenzen 22 sowie die Schwellwerte 23 bereitgestellt wurden:

5 (----- Appendix 2 -----)

Insbesondere hinsichtlich der Parameter „matrix_sizes“, „submatrix_sizes“, „sweep_widths“, „ppm_offsets“, „nucleus_species“ sowie „mask_lower_threshold“ ist nun jeweils der numerische Wertebereich festgelegt.

10

Sobald die Suchbereichsgrenzen 22 sowie die Schwellwerte 23 festgelegt sind, kann die Mustererkennungs-Software 20 mit der eigentlichen Mustererkennung beginnen. Hierzu werden die aufgezeichneten 2D-Spektren „2rr“ sowie die 3D-Spektren „3rrr“ mit den Suchbereichen der Kreuzsignalmuster-Suchmaske (bzw: einer ihrer Unter-Suchmasken) korreliert, um eine

15 Bewertungszahl für das Vorliegen des durch die Kreuzsignalmuster-Suchmaske erfassten Kreuzsignalmusters zu erhalten.

Bei dem in Fig. 4 gezeigten Beispiel werden die Werte des Spektrums, die innerhalb der Signalsuchbereiche 9, 11, 13, 15, 17, 19 liegen, aufsummiert, um so die Bewertungszahl für das 20 Vorliegen des Kreuzsignalmusters zu erhalten. Wenn die erwarteten Signalpeaks innerhalb der Signalsuchbereiche 9, 11, 13, 15, 17, 19 auftreten, dann ergibt sich eine hohe Bewertungszahl für das gesuchte Kreuzsignalmuster. Dies bedeutet, dass das gesuchte Kreuzsignalmuster mit hoher Wahrscheinlichkeit vorliegt. Wenn die erwarteten Peaks in den Signalsuchbereichen 9, 11, 13, 15, 17, 19 dagegen ganz oder teilweise fehlen, dann ergibt sich eine dementsprechend 25 niedrige Bewertungszahl. In diesem Fall ist es unwahrscheinlich, dass das gesuchte Kreuzsignalmuster vorliegt.

Zur Berechnung der Bewertungszahl wird ein sog. Masken-Scan 24 durchgeführt, bei dem die Koordinaten der chemischen Verschiebung in den verschiedenen Raumrichtungen nacheinander hochgezählt werden, um so das gesamte Spektrum abzurastern. Der so erzeugte Koordinatenwert wird mit den Maskendaten 25 verglichen. Wenn der Koordinatenwert außerhalb von sämtlichen Signalsuchbereichen der Kreuzsignalmuster-Suchmasken liegt, so bleibt die Bewertungszahl unverändert. Wenn der neu generierte Koordinatenwert dagegen innerhalb

eines Suchbereichs liegt, so wird der zu diesem Koordinatenwert gehörige Spektrenwert zu der Bewertungszahl addiert. Mit Hilfe eines derartigen Masken-Scans 24 lässt sich die Bewertungszahl für eine bestimmte Kreuzsignalmuster-Suchmaske bzw. für eine bestimmte Unter-Suchmaske der Kreuzsignalmuster-Suchmaske schnell und einfach ermitteln.

5

Bisher war davon ausgegangen worden, dass die Kreuzsignalmuster-Suchmaske lediglich eine Anzahl von Signalsuchbereichen umfasst, wobei innerhalb der Suchbereichsgrenzen das Auftreten von Signalpeaks erwartet wird. Entsprechend können aber auch Leerbereiche definiert werden, wobei das Auftreten eines Signalpeaks innerhalb der Leerbereichs-Grenzen zu einer

10 Verringerung der Bewertungszahl führt. Derartige Leerbereiche stellen also gewissermaßen „verbotene“ Bereiche dar, in denen keine Signalpeaks auftreten dürfen.

Zur Verbesserung der Erkennungsgenauigkeit kann versucht werden, die Qualität der 2D- bzw. 3D-Spektren vor der Durchführung der Mustererkennung mittels einer Faltungsoperation zu verbessern. Hierzu kann ein ideales Gauß-Signal, dessen Größe und Ausdehnung in etwa der Größe und Ausdehnung der erwarteten NMR-Signalpeaks entspricht, mit dem Spektrum gefaltet werden. Auf diese Weise können Artefakte unterdrückt und verschmierte Peaks besser aufgelöst werden.

20 In Fig. 5A ist dargestellt, dass die Auswertung von zwei aminosäuretyp-spezifischen 2D-Spektren sowie von einem Paar von Tripelresonanzexperimenten vier Teilergebnisse 26, 27, 28, 29 liefert. Zur Auswertung der zwei aminosäuretyp-spezifischen 2D-Spektren können zwei hierfür geeignete Unter-Suchmasken vorgesehen sein, und zur Auswertung der zwei Tripelresonanzexperimente können zwei weitere Unter-Suchmasken der Kreuzsignalmuster-Suchmaske definiert sein. Jede Unter-Suchmaske der betrachteten Kreuzsignalmuster-Suchmaske erzeugt ein Teilergebnis 26, 27, 28, 29. Beispielsweise enthält ein Teilergebnis für ein bestimmtes 2D-Spektrum die chemischen Verschiebungen der innerhalb der Signalsuchbereiche aufgefundenen Peaks zusammen mit der Bewertungszahl für die Unter-Suchmaske.

25 26, 27, 28, 29, die bei der Auswertung der zwei aminosäuretyp-spezifischen 2D-Spektren sowie des Paars von Tripelresonanzexperimenten entstehen, werden der Merging-Einheit 30 zugeführt. Aufgabe der Merging-Einheit 30 ist es, die unterschiedlichen Teilergebnisse zu einer Ergebnisliste 31 zusammenzufassen. Hierzu werden die

30 31 zusammenzufassen. Hierzu werden die

chemischen Verschiebungen in den Überlappbereichen der einzelnen Teilergebnisse verglichen. Insbesondere auf der Grundlage der chemischen Verschiebungen der Kerne H_N sowie N kann dann über ein Intervall von $H_N \pm \Delta H_N$ sowie von $N \pm \Delta N$ das Merging, also eine Zusammenfassung der Teilergebnisse zur Ergebnisliste 31, vorgenommen werden. Ein Eintrag in der Ergebnisliste 31 umfasst einen sog. Shift-Vektor, in dem die innerhalb von zwei aufeinanderfolgenden Aminosäuren auftretenden chemischen Verschiebungen aufgelistet sind, sowie die insgesamt für das Vorliegen der Zweiergruppe von Aminosäuren ermittelte

5 Bewertungszahl.

10 Im Anschluß wird die so entstandene Ergebnisliste gewichtet. Der Wichtungsvorgang wird von der Cleaning-Einheit 32 vorgenommen. Die Cleaning-Einheit 32 überprüft die Plausibilität des gefundenen Ergebnisses, indem sie mit einer Wichtungsfunktion die Shift-Vektoren auf ihre Richtigkeit überprüft.

15 Im folgenden ist die Ergebnisliste einer Kreuzsignalmuster-Suchmaske aufgeführt, welche nach den von dem Aminosäurepaar N-S hervorgerufenen Kreuzsignalmuster sucht. Wie aufgrund der durch das Programm „selma“ durchgeführten Fragmentierung der Primärsequenz nicht anders zu erwarten, wurde lediglich ein Eintrag gefunden, da die Zweiergruppe von aufeinanderfolgenden Aminosäuren N-S in der Primärsequenz des betrachteten Proteins 20 OPR nur einmal vorkommt. Der Eintrag in die Ergebnisliste enthält eine Auflistung von chemischen Verschiebungen sowie eine Bewertungszahl, die unter der Rubrik „Resp“ aufgeführt ist.

(----- Appendix 3 -----)

25 Zur Überprüfung der Plausibilität des für das Aminosäurepaar N-S aufgefundenen Ergebnisses wird außerdem eine Mustersuche in den für die komplette Proteinhauptkette aufgezeichneten 3D-Experimenten durchgeführt. Hierbei erhält man eine umfangreichere Ergebnisliste, deren Einträge wiederum eine Anzahl von chemischen Verschiebungen sowie eine Bewertungszahl umfassen:

(----- Appendix 4 -----)

Anhand der übereinstimmenden chemischen Verschiebungen kann erkannt werden, dass es sich bei dem Eintrag #53 um den Eintrag für das Aminosäurepaar N-S handelt. Dies bestätigt die Konsistenz der bei den beiden Mustersuchen aufgefundenen Ergebnisse.

5 Nachdem für die einzelnen in der Proteinsequenz enthaltenen Aminosäurepaare Teilzuordnungen vorgenommen worden sind, müssen diese Teilzuordnungen auf die Primärsequenz abgebildet werden. Dieser Schritt wird als Sequenz-Mapping 33 bezeichnet. Auf der Grundlage der Primärsequenz werden die Aminosäurepaare in den Ergebnislisten gesucht und mit der höchsten Wichtung beginnend auf die Sequenz abgebildet. Nach jeder Iteration wird die Verkettung der einzelnen Paare überprüft und es werden so Fragmente von 2, 3, 4 etc. Aminosäuren gebildet. Nach Abschluß der Iterationsroutine werden die fehlenden Fragmente in den Ergebnislisten der Mustersuche für die 3D-Experimentepaare gesucht und auf die Sequenz abgebildet. Nach Abschluß der Hauptkettensuche erfolgt eine gezielte Suche in den Ergebnislisten aller Seitenkettenexperimente.

15

Als Ergebnis des Sequenz-Mappings 33 erhält man eine vollständige sequenzielle Zuordnung 34 der in den verschiedenen Spektren auftretenden Signalpeaks zu den verschiedenen Aminosäuren der Sequenz. Damit ist die Aufgabe, eine möglichst automatisierbare Zuordnung der Spektralpeaks durchzuführen, ausgeführt.

20

Diese Zuordnung sowie die für die verschiedenen magnetisch aktiven Kerne des Proteins aufgefundenen chemischen Verschiebungen können als Ausgangspunkt für die automatische Zuordnung von NOESY-Spektren 36 herangezogen werden. Diese Zuordnung der NOESY-Spektren 36 zu den verschiedenen magnetisch aktiven Kernen des Proteins wird durch das

25

Programm ARIA 35 durchgeführt. Die Bezeichnung ARIA steht für „Ambiguous Restraints for Iterative Assignment“. Der Startpunkt für ARIA ist eine nahezu vollständige Zuordnung der protonenchemischen Verschiebungen, welche zusammen mit einer Liste der Amplituden der NOE-Kreuzsignale und ihrer Frequenzkoordinaten einem Strukturberechnungs-Programm (im speziellen Fall „Explor“) übergeben wird. Die zentrale Aufgabe des Programms ARIA 35 ist die Zuordnung der NOE während der Strukturrechnung durch die Annahme einer mehrfachen Bedeutung der einzelnen Kreuzsignale und die Anwendung einer iterativen Zuordnungsstrategie für dieselben.

Patentansprüche

1. Analysesystem zur automatisierten Analyse eines Satzes von NMR-Spektren, welcher für
5 eine n Aminosäuren umfassende Polypeptidkette aufgezeichnet wurde, mit
einer Bibliothek von Kreuzsignalmuster-Suchmasken, wobei eine Kreuzsignalmuster-
Suchmaske zur spezifischen Erfassung der NMR-Signale eines Fragments der untersuchten
Polypeptidkette vorgesehen ist,
Mitteln zur Auswahl der zur Analyse benötigten Kreuzsignalmuster-Suchmasken aus der Bi-
10 bliotheke von Kreuzsignalmuster-Suchmasken entsprechend der Primärsequenz der Polypep-
tidkette, welche für jedes in der Primärsequenz enthaltene Fragment die zugehörige Kreuzsi-
gnalmuster-Suchmaske auswählen,
Mitteln zur Mustererkennung, welche die verschiedenen Ergebnisse der entsprechend der
Primärsequenz der Polypeptidkette ausgewählten Kreuzsignalmuster-Suchmasken verbinden
15 und mit dem Satz von NMR-Spektren korrelieren,
Mitteln zur Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der Polypep-
tidkette entsprechend dem Ergebnis der Mustererkennung.
2. Analysesystem nach Anspruch 1, dadurch gekennzeichnet, dass
20 die Fragmente der untersuchten Polypeptidkette jeweils zwei bzw. drei bestimmte aufeinan-
derfolgende Aminosäuren umfassen.
3. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass
der Satz von NMR-Spektren sowohl NMR-Experimente zur Analyse der Hauptkettensignale
25 als auch NMR-Experimente zur Analyse der Seitenkettensignale umfasst.
4. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass
die zur Analyse der Hauptkettensignale verwendeten NMR-Experimente 3D-Experimente,
und insbesondere 3D-Experimenten der Typen CBCA(CO)NNH, CBCANNH,
30 HA(CO)NNH, HANNH, HAHB(CO)NNH, HAHBNNH, HN(CA)CO, HNCO,
HN(CO)CA, HNCA umfassen.
5. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass

die zur Analyse der Seitenkettensignale verwendeten NMR-Experimente umfassen.

6. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass

5 die zur Analyse der Haupt- und Seitenkettensignale verwendeten NMR-Experimente aminosäuretyp-spezifische 2D-Experimente umfassen, welche selektiv sind für einen Aminosäuretyp oder für eine Gruppe von Aminosäuretypen.

7. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass

10 die zur Analyse der Haupt- und Seitenkettensignale erforderlichen aminosäuretyp-spezifischen 2D-Experimente entsprechend der Primärsequenz der Polypeptidkette festgelegt werden.

8. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass

15 die zur Analyse der Haupt- und Seitenkettensignale verwendeten NMR-Experimente eine Kombination von 2D- und 3D-Experimenten umfassen.

9. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass ausgehend von der Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der

20 Polypeptidkette eine Auswertung der NMR-Spektren zur Ermittlung von chemischen Verschiebungen und Kopplungskonstanten erfolgt.

10. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass ausgehend von der Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der

25 Polypeptidkette die chemischen Verschiebungen zusammengefasst und auf ihre Richtigkeit überprüft werden.

11. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass der Satz von NMR-Spektren Spektren des Typs NOESY umfasst, deren Auswertung insbesondere Informationen über die Abstände der verschiedenen Kerne der Polypeptidkette von-

30 einander liefert.

12. Analysesystem nach Anspruch 11, dadurch gekennzeichnet, dass

die Zuordnung der NMR-Spektren des Typs NOESY zu den verschiedenen Kernen der Polypeptidkette anhand der für die Kerne ermittelten chemischen Verschiebungen erfolgt.

13. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die bei der Auswertung der NMR-Spektren erhaltenen Größen als Eingangsgrößen für Strukturberechnungs-Programme dienen.
14. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Kreuzsignalmuster-Suchmasken jeweils eine Anzahl von vordefinierten Signalsuchbereichen umfassen, wobei durch das Auftreten von NMR-Signalen innerhalb der Bereichsgrenzen eines Signalsuchbereichs die Wahrscheinlichkeit erhöht wird, dass das durch die Kreuzsignalmuster-Suchmaske definierte Signalmuster vorliegt.
15. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Kreuzsignalmuster-Suchmasken jeweils eine Anzahl von vordefinierten Leerbereichen umfassen, wobei durch das Fehlen von NMR-Signalen innerhalb der Bereichsgrenzen eines Leerbereichs die Wahrscheinlichkeit erhöht wird, dass das durch die Kreuzsignalmuster-Suchmaske definierte Signalmuster vorliegt.
16. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass ausgehend von der erwarteten Anzahl der NMR-Signale in den Spektren durch Iteration die Schwellwerte und Suchbereiche für die Kreuzsignalmuster-Suchmasken ermittelt werden.
17. Analysesystem nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass eine Kreuzsignalmuster-Suchmaske eine Vielzahl von Unter-Suchmasken zur Analyse der verschiedenen NMR-Spektren des aufgezeichneten Satzes von NMR-Spektren umfassen kann.
18. Verfahren zur automatisierten Analyse eines Satzes von NMR-Spektren, welcher für eine n Aminosäuren umfassende Polypeptidkette aufgezeichnet wurde, gekennzeichnet durch folgende Schritte:
 - a) Auswählen der zur Analyse benötigten Kreuzsignalmuster-Suchmasken aus einer Bibliothek von Kreuzsignalmuster-Suchmasken, wobei eine Kreuzsignalmuster-Suchmaske zur spe-

zifischen Erfassung der NMR-Signale eines Fragments der untersuchten Polypeptidkette vorgesehen ist, und wobei das Auswählen der benötigten Kreuzsignalmuster-Suchmasken entsprechend den in der Primärsequenz enthaltenen Fragmenten erfolgt;

b) Durchführen einer Mustererkennung durch Korrelieren der verschiedenen ausgewählten

5 Kreuzsignalmuster-Suchmasken mit dem Satz von NMR-Spektren;

c) Zuordnen der NMR-Signale zu den verschiedenen Spinsystemen der Polypeptidkette entsprechend dem Ergebnis der in Schritt b) durchgeführten Mustererkennung.

19. Verfahren nach Anspruch 18, dadurch gekennzeichnet, dass

10 die Fragmente der untersuchten Polypeptidkette jeweils zwei bzw. drei bestimmte aufeinanderfolgende Aminosäuren umfassen.

20. Verfahren nach Anspruch 18 oder Anspruch 19, dadurch gekennzeichnet, dass der Satz von NMR-Spektren sowohl NMR-Experimente zur Analyse der Hauptkettensignale 15 als auch NMR-Experimente zur Analyse der Seitenkettensignale umfasst.

21. Verfahren nach einem der Ansprüche 18 bis 20, dadurch gekennzeichnet, dass die zur Analyse der Hauptkettensignale verwendeten NMR-Experimente 3D-Experimente, und insbesondere 3D-Experimenten der Typen CBCA(CO)NNH, CBCANNH,

20 HA(CO)NNH, HANNH, HAHB(CO)NNH, HAHBNNH, HN(CA)CO, HNCO, HN(CO)CA, HNCA umfassen.

22. Verfahren nach einem der Ansprüche 18 bis 21, dadurch gekennzeichnet, dass die zur Analyse der Seitenkettensignale verwendeten NMR-Experimente Experimente der 25 Typen HCCH-COSY, HCCH-TOCSY, HCC(CO)NH-TOCSY umfassen.

23. Verfahren nach einem der Ansprüche 18 bis 22, dadurch gekennzeichnet, dass die zur Analyse der Haupt- und Seitenkettensignale verwendeten NMR-Experimente aminosäuretyp-spezifische 2D-Experimente umfassen, welche selektiv sind für einen Aminosäure-

30 typ oder für eine Gruppe von Aminosäuretypen.

24. Verfahren nach einem der Ansprüche 18 bis 23, dadurch gekennzeichnet, dass

die zur Analyse der Haupt- und Seitenkettensignale erforderlichen aminosäuretyp-spezifischen 2D-Experimente entsprechend der Primärsequenz der Polypeptidkette festgelegt werden.

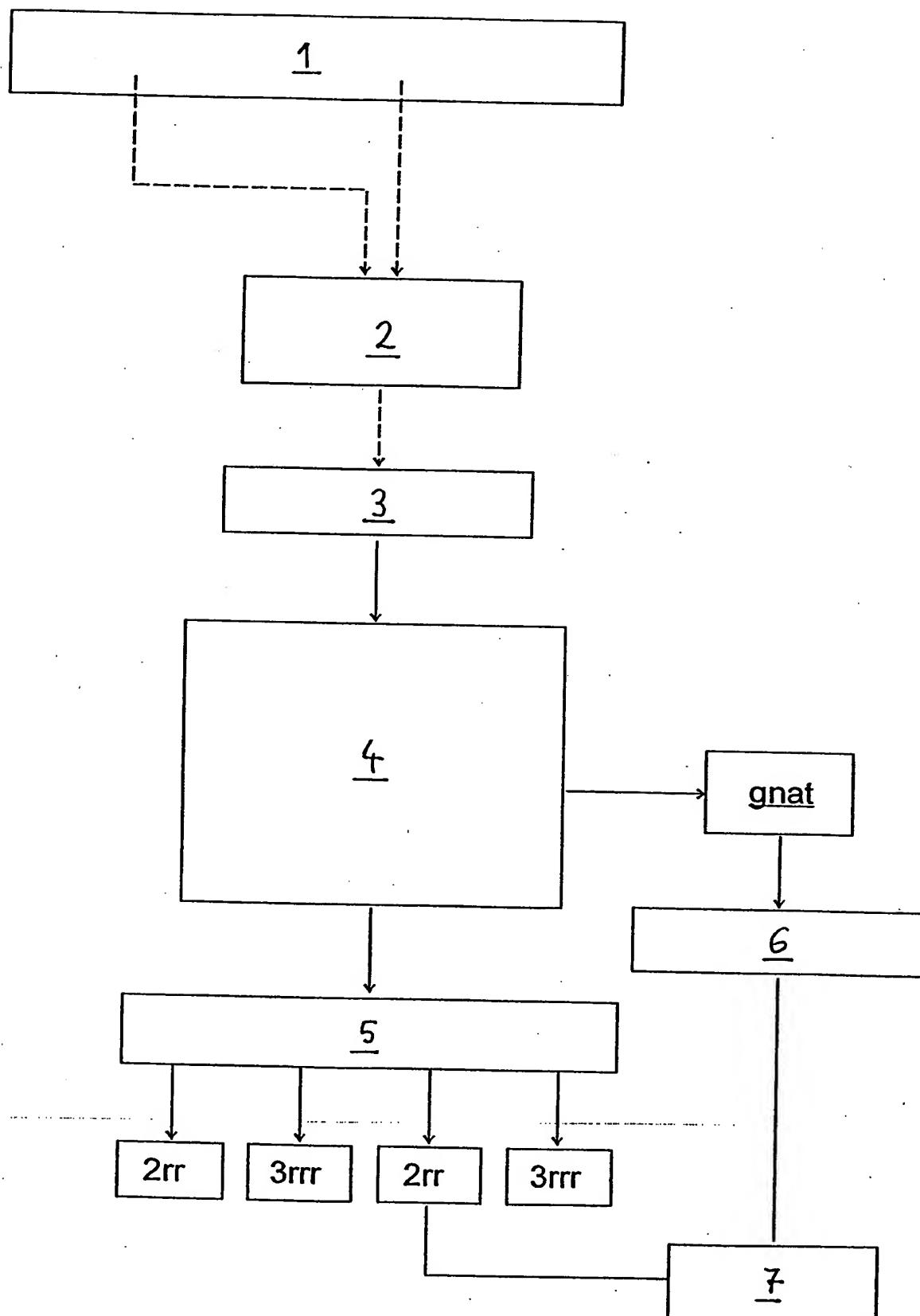
- 5 25. Verfahren nach einem der Ansprüche 18 bis 24, dadurch gekennzeichnet, dass zur Analyse der Haupt- und Seitenkettensignale eine Kombination von 2D- und 3D-Experimenten verwendet wird.
- 10 26. Verfahren nach einem der Ansprüche 18 bis 25, dadurch gekennzeichnet, dass ausgehend von der Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der Polypeptidkette eine Auswertung der NMR-Spektren zur Ermittlung von chemischen Verschiebungen und Kopplungskonstanten durchgeführt wird.
- 15 27. Verfahren nach einem der Ansprüche 18 bis 26, dadurch gekennzeichnet, dass ausgehend von der Zuordnung der NMR-Signale zu den verschiedenen Spinsystemen der Polypeptidkette die chemischen Verschiebungen zusammengefasst und auf ihre Richtigkeit überprüft werden.
- 20 28. Verfahren nach einem der Ansprüche 18 bis 27, dadurch gekennzeichnet, dass der Satz von NMR-Spektren Spektren des Typs NOESY umfasst, und dass die Zuordnung der NMR-Spektren des Typs NOESY zu den verschiedenen Kernen der Polypeptidkette ausgehend von den für die verschiedenen Kerne ermittelten chemischen Verschiebungen durchgeführt wird.
- 25 29. Verfahren nach einem der Ansprüche 18 bis 28, dadurch gekennzeichnet, dass die bei der Auswertung der NMR-Spektren erhaltenen Größen als Eingangsgrößen für Strukturberechnungs-Programme verwendet werden.

Zusammenfassung**Vorrichtung und Verfahren zu Zuordnung der NMR-Signale von Polypeptiden**

5

Die NMR-Spektroskopie wird zunehmend für die Bestimmung der Struktur von Proteinen eingesetzt. Erfindungsgemäß werden zur gemeinsamen Auswertung der 3D-Tripelresonanzexperimente sowie der aminosäuretyp-spezifischen 2D-Spektren Kreuzsignal-
10 muster-Suchmasken verwendet. Mithilfe der erfindungsgemäßen Kreuzsignalmuster-
Suchmasken lassen sich mittels Mustererkennung sämtliche von einer oder einer Gruppe von
Aminosäuren, deren genaue Abfolge vorher aus der Proteinsequenz abgeleitet wurde und die
Grundlage für die spezifische Mustererkennung bietet, verursachten Signalpeaks gemeinsam
15 erfassen und auswerten. Die Erfindung ermöglicht eine vollautomatische Zuordnung der in
den verschiedenen Spektren vorkommenden Signalpeaks zu den magnetisch aktiven Kernen
des Proteins.

Z U S A M M E N F A S S U N G



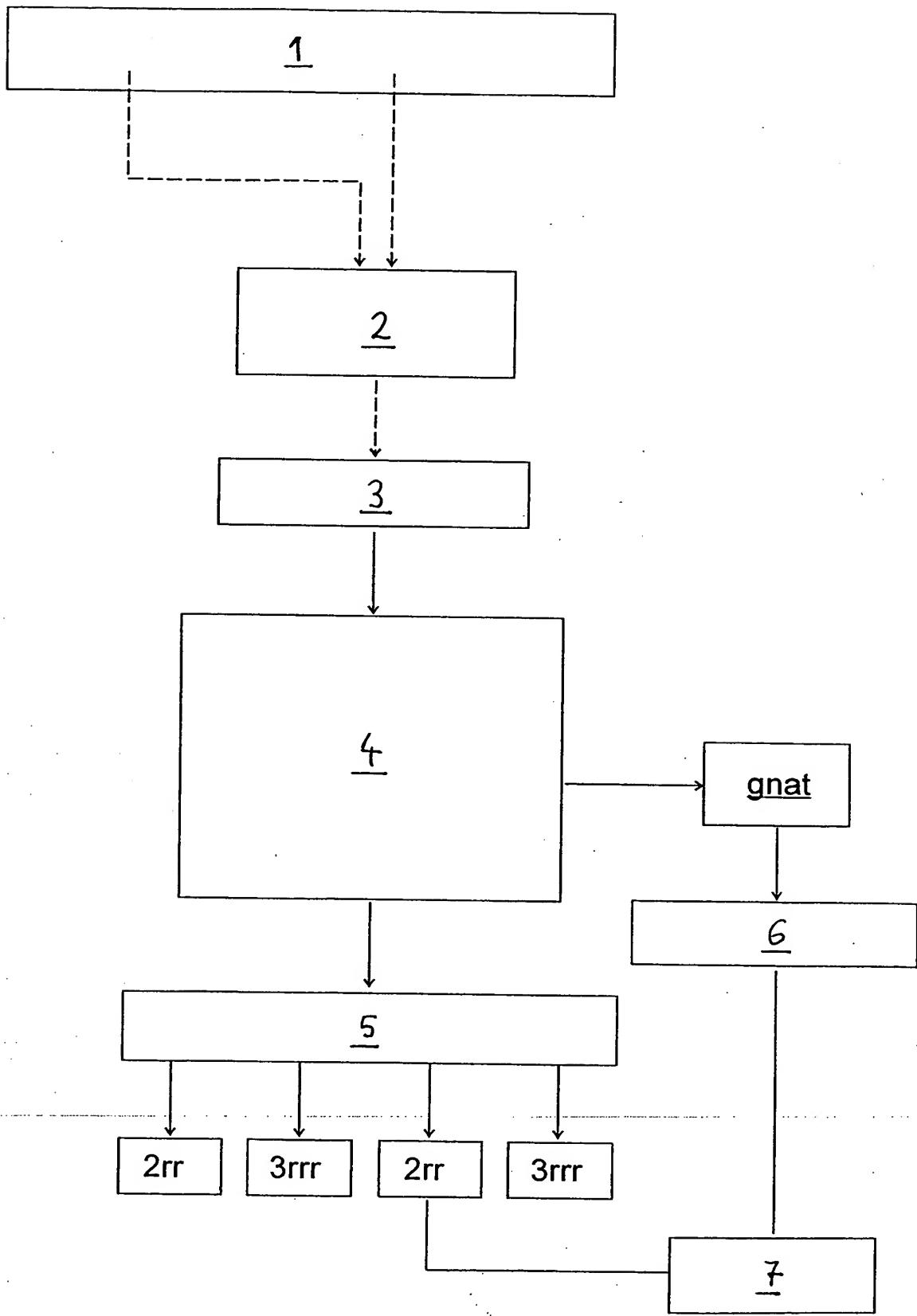


Fig. 1

	A	D	E	G	L	N	Q	R	S	T	W	P	I	V	F	H	Y	C	K	M
A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
D	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
E	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
G	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	
L	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
N	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Q	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
R	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
T	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
W	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
P	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
V	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
F	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
H	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Y	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
K	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
M	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

Fig. 2

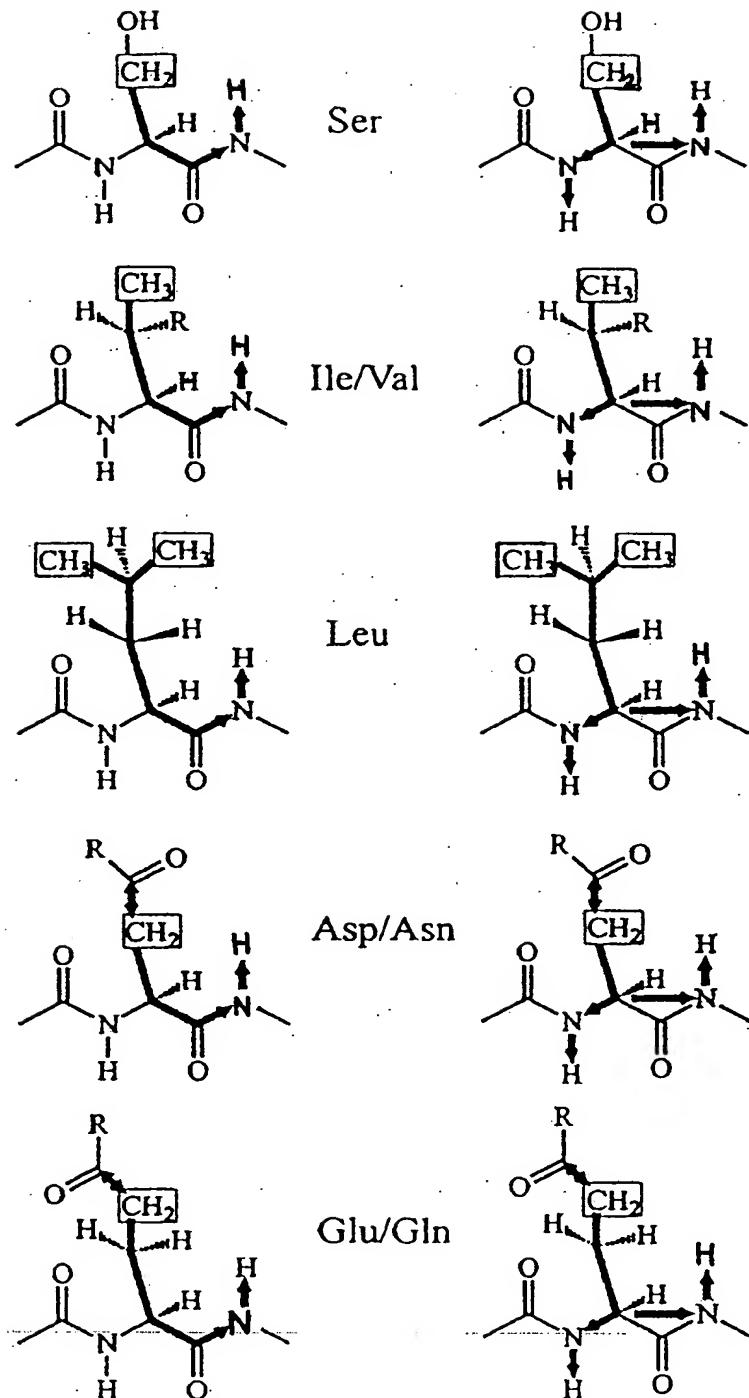


Fig. 3

4/6

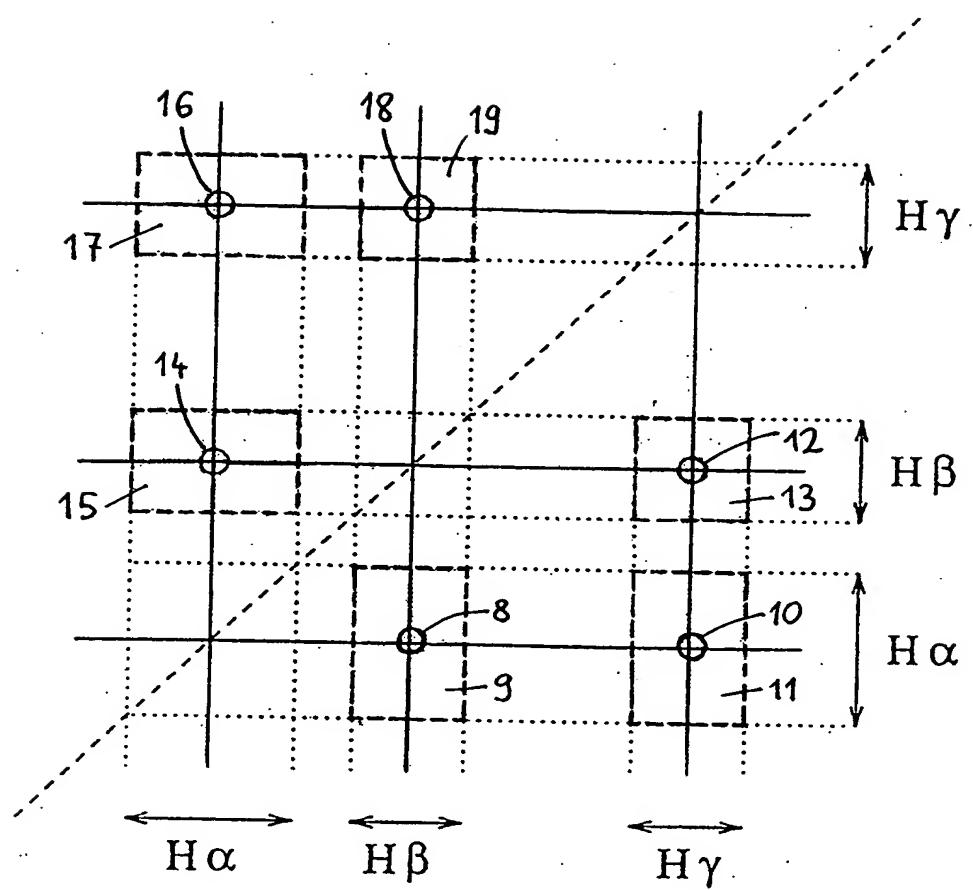


Fig. 4

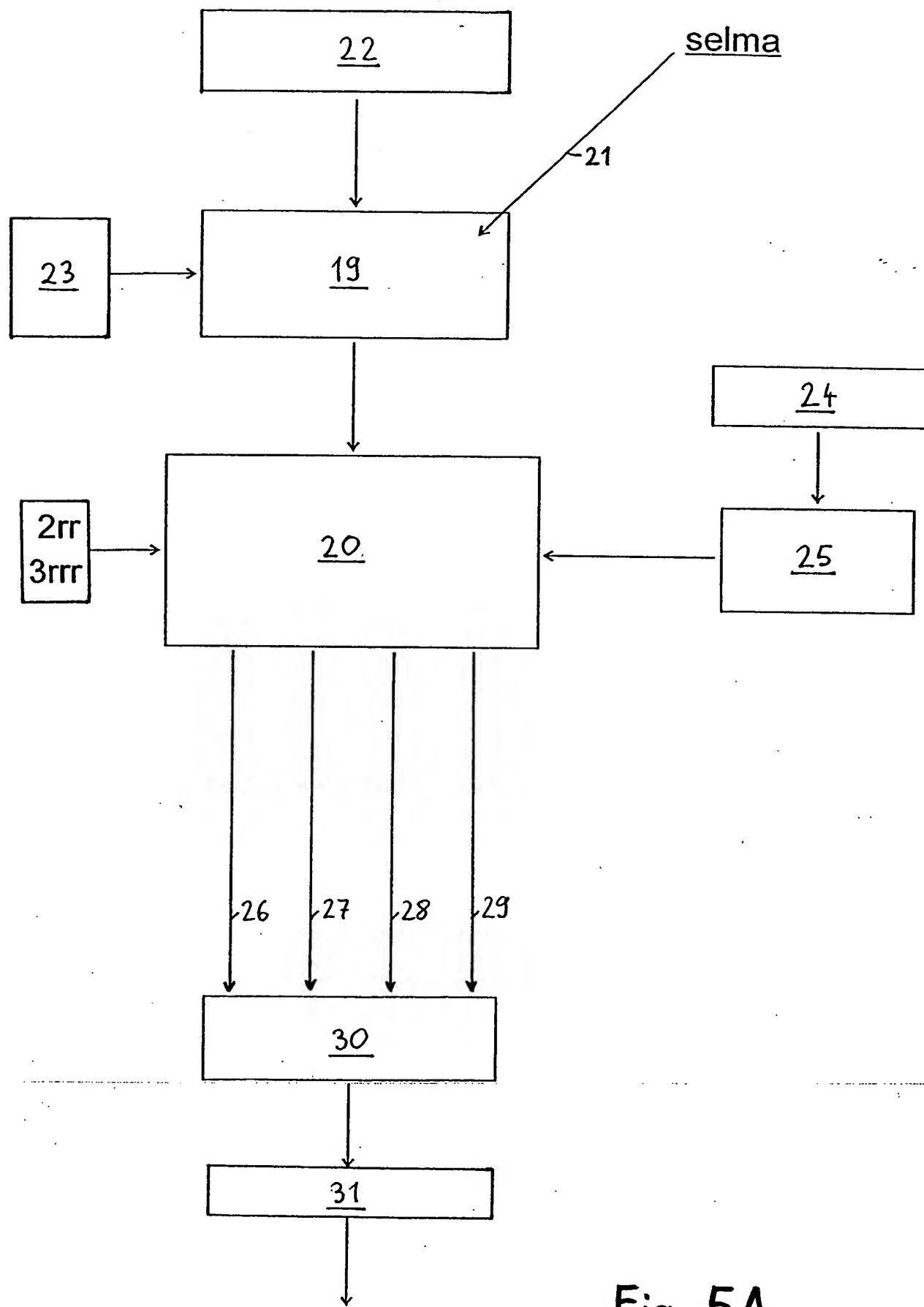


Fig. 5A

6/6

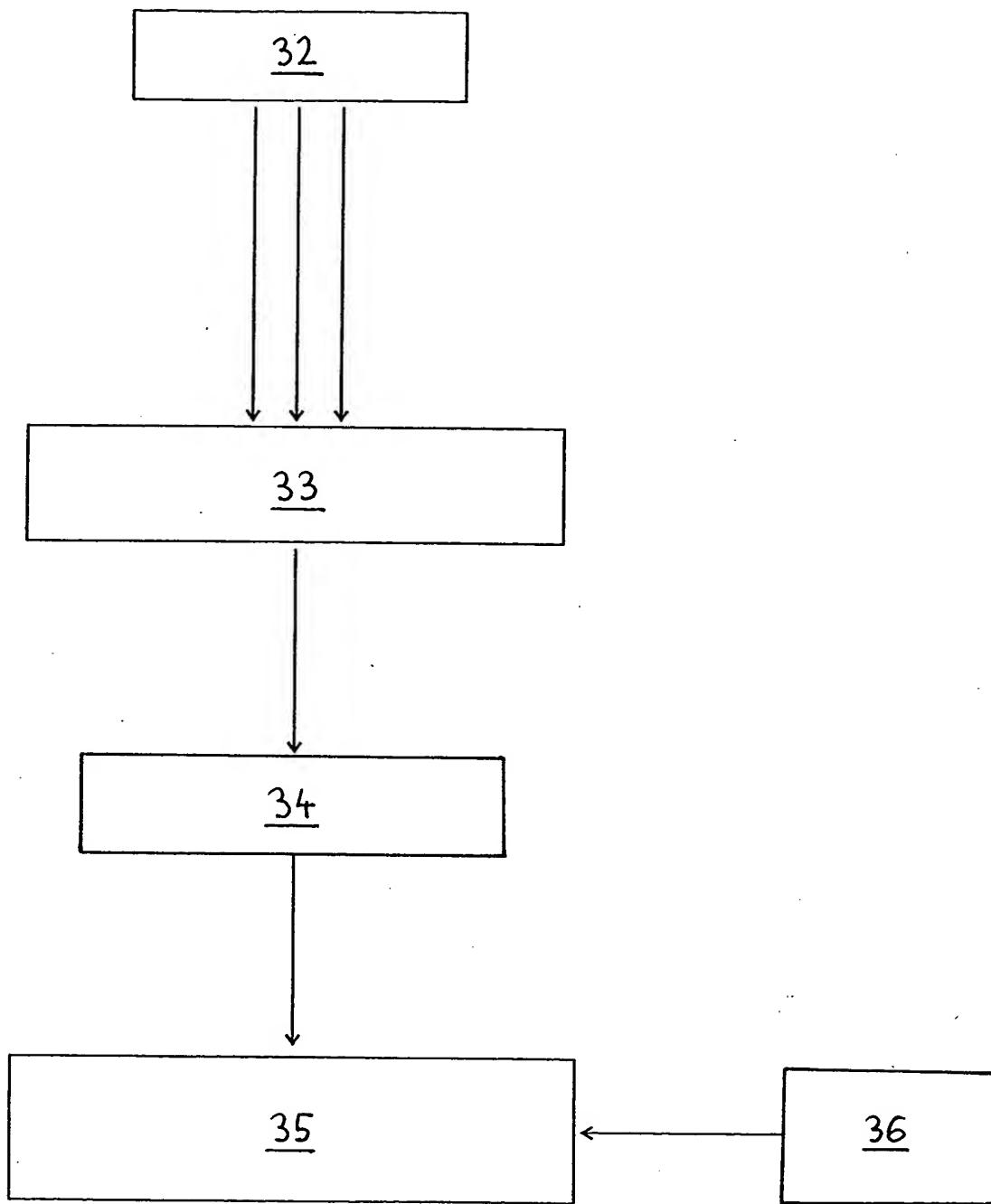


Fig. 5B